

データベース計画

関根 哲

日本電子株式会社

〒196 昭島市武蔵野3-1-2

1. はじめに

表面分析の歴史は20余年と短い、現在、表面科学の研究、材料開発、産業界に置く品質評価などの道具として巾広く用いられている。現在ではデータの測定、処理、解析、報告書の作成など、分析の全ての過程がコンピュータ化されている。しかし、コンピュータ化されてからそう長い時間は経っておらず、まだまだ満足できる状況ではない。更に分析精度をあげるため、分析結果の普遍性を高めるため、そして便利さを追求するために多くの改善が望まれている。

現状では、同じ試料を測定しても、使用する装置、測定する人、装置の調整方法などが違えば異なるスペクトルが得られることが多々あり、そのどれが正しいスペクトルなのか判断することは容易ではない。スペクトルの解析段階でも同様である。同一スペクトルを解析しても、解析方法（ピーク強度やピーク位置の決め方、補正方法など）、解析に用いられる標準データ、物理化学パラメータなどが異なれば、違った結果が出てしまうことがある。これについてもどれが正しいのかを判断することはなかなか容易ではない。

この問題を解決するための3つのキーワードは、標準化、データベース、共通データ処理環境であろう。標準化された方法を用い、必要なデータはデータベースから引出し、これを用いて解析を行う。また、共通データフォーマットにアクセスする共通データ処理環境を用い種々の解析方法を研究できる。このような状況が実現できれば、表面電子分光分析はさらに信頼性の高いものとなろう。ここでは表面分析研究会（旧VAMAS-SCA-JAPAN）で手掛けているデータベース計画について述べる。

2. 現存するデータベースとデータ処理ソフトウェア

表面分析関係のデータベースやデータ処理ソフトウェアはあまり多くない。筆者の知る限りを列挙する。

NIST XPS DATABASE¹⁾: アメリカのNational Institute of Standards and Technology (NIST) が作成したフロッピーベース (MS-DOS) のデータベースで、XPSとAESのピークエネルギー値が納められている。条件検索により必要な物質のピークエネルギー値を選び出せる。

SSS: アメリカ真空学会 (AVS) が発行しているスペクトルを掲載した雑誌 Surface Science Spectra (SSS) で、1992年はじめに第1号が発行された。印刷ベースでスタートするが、後日デジタル化する計画と聞いているが、まだ実現していない。

COMPRO/データバンク: 当表面分析研究会が開発したCommon Data Processing System (COMPRO, 共通データ処理環境) は表面分析のためのデータ処理プラットフォームで、約400本のスペクトルデータバンクと物理化学パラメータデータバンクが内蔵されている。フロッピーベース (MS-Windows) で供給される。

Seah: イギリスのNPLのM. Seahらの開発したスペクトルデータベースで、5kV

と10kVで測定した高分解能(0.25eVステップ)の積分スペクトルが納められている。強度軸は透過関数補正が施された絶対基準で表されている。フロッピーベース(MS-DOS)で、近々供給開始の見込み。

Tougaard²⁾：デンマークのオーデンス大学のTougaard教授の開発したXPS/AESのスペクトル解析ソフトウェアパッケージ。非弾性ピーク形状のもたらず情報から、膜の島状構造、表面近傍の深さ方向組成分布などの定量解析を行う。Quantitative Analysis of Surface Electron Spectra (QUASES)という名称。

3. 表面分析研究会の取り組み

データベースの検討は1990年の初め頃から始まった。その動機は、一つには共通データ処理環境を開発過程で、それを有効に活用するためにはどうしても標準スペクトルデータなどを納めたデータベースが必要なことを痛感したこと、また、当時アメリカのAVSのグループが壮大なデータベースを構築しようとしていたこと知って刺激されたことである。当研究会としては、どの様なデータベースを構築するか喧々ごうごう議論した。色々な形態のデータベースが考えられるが、測定条件を明記したスペクトルを集めたスペクトルデータベースは特に重要との認識を得た。そして結果としてつぎの二つのタイプに取り組んだ。

電話回線データベースの試作³⁾：

小規模なネットワークベースのデータベースをつくり、経験することによって問題を検討する目的で、パソコンをホスト局にしたデータベースを試作した。市販のホスト局のソフトを用いて、ボード上にスペクトルリストを置き、ダウンロード、アップロードができるようにした。このシステムは極めて短期間ではあったが、約50本のスペクトルを収納し実際に稼動した。

COMPRO/データバンク：

電話回線データベース用に集められたスペクトルの他に、その後集めたスペクトルを加え、400本のスペクトルデータバンクがCOMPROに組み込まれた。これらのスペクトルの中には、後藤教授の標準CMAで測定された絶対基準のスペクトルも格納されている。さらに、物理化学パラメータデータバンクも追加され、現在に至っている。

省際ネット⁴⁾：

科学技術庁では1994年より省際ネットワークの整備・運用に係わる基盤技術に関する研究が開始された。この研究は、つくば、東京、関西を結ぶ新たなネットワークを構築し、それにいくつかのデータベースを接続し、各機関がネットワークを利用して相互にデータを利用しあうことを目的としている。このネットワークに組み込まれるデータベースとして表面分析データベースが採用された。表面分析データベースは、①スペクトルデータベース、②スペクトル解析のための物理パラメータデータベースの2種類が含まれる。当研究会では、この機会を利用して5年がかりで表面分析データベースの構築を行うことを決定した。

4. 計画の現状

コンピュータ環境：

データベースの主要部分はスペクトルデータベースで、少なくとも5000本以上は収納したい。サーバタイプのホストコンピュータHP-9000/E55を金属材料技術研究所に設置し、これをインターネットに接続する。ホストコンピュータとのアクセスにはイ

インターネットの他に電話回線 (MODEM) も併設される計画である。データベースの基本ソフトウェアはオラクル社のリレーショナル・データベース管理ソフトウェアを使用する。ユーザーのパソコン (クライアント PC) とホストコンピュータ (サーバシステム) とのコミュニケーションは NTT 社の VGUIDE というソフトウェア (ミドルウェア) を介して行う。このミドルウェアの特徴は Visual Basic や Excel など市販のソフトウェアとリンクして使用できることである。当研究会で開発した共通データ処理システムは Visual Basic で書かれているので、このデータベースは共通データ処理システムとリンクして使用できる。

データの収集:

現在は、データベースの中に納めるデータを、研究会のメンバーで手分けして測定中である。当研究会の中には、電子材料、金属材料、有機材料、無機材料の4つの分科会があり、材料の種類別に材料の手配、測定条件の検討、必要な情報の洗いだし、データの提出の催促などに当たっている。データベース用の質の高い (検証可能性の高い) データを測定するためには、いろいろな試行錯誤が予想され、研究を行いながら進めて行くことになる。むしろここから新たな研究テーマが生まれてくることが期待される。

データの転送と変換:

測定したデータをデータベースへ納めるためには、まず測定した装置のコンピュータからデータを PC へとりだし、それぞれのデータフォーマットから共通の VAMAS フォーマットへ変換する必要がある。現在、各メーカーからフォーマット情報の提供を受け、変換ソフトウェアを開発中である。

VAMAS フォーマットはデータ転送用に開発された、最少限のパラメータで構成される小さいフォーマットである。データベースに対する想定はされていなかったもので、データベース用に使用するためには足りないパラメータがある。ISO の標準フォーマットとして提案されているものは、このような背景から、VAMAS フォーマットに日本が提案した Standard Data Information Format (データベースを意識して最少限追加すべきパラメータをセットにしたもの) を挿入した形式となっている。

5. スペクトルデータベースの概念設計

コンピュータと基本ソフトウェアは購入品であるが、その利用形態の開発は我われ表面分析研究会が創造性を発揮できる部分である。逆に、ここに創造性を発揮しなければ、意味の小さいものになってしまう。

データ検索

本データベースはリレーショナルであり、キーワードを用いて必要とするデータの組を選び出すことが可能である。適切なキーワードの選択を行なう必要がある。これに関連して検索リストの表示には、一行のタイトルは必要ではなからうか。データが選択されてもファイル名のみでは中身がうかがい知れない。

データの構造

スペクトルデータは VAMAS フォーマットで交換されるが、データベースの中はどんな形式が良いかは別である。第1に VAMAS フォーマットは ASCII 形式であり、データサイズが大きい。内部格納には BINARY の方がコンパクトが良い。第2に指定されたデータが単一のスペクトルのことも有れば、複数のスペクトルのセットのこともある。*

データの格納ではこの点に配慮が必要となるであろう。第3に提案中の ISO フォーマ

ットでも足りないパラメータが現れて来るであろう。このような場合、補助ファイルが必要となるであろう。データ格納にはこの点を考慮しておくべきと思う。

*) 例えば、純元素のワイドスペクトルとナロースペクトル、イオンスパッタによる変化過程、還元過程など。

保守の方法

スペクトルは時間と共に増えるので、新しいスペクトルを加えたり、不要となったスペクトルを削除したり、キーワードの変更、あるいはデータフォーマットの内容に変更を加えたりした時など保守法についての検討をしておかなければならない。

スペクトルデータベースに登録して利用されるデータは、エネルギー軸、強度軸、エネルギー分解能、S/Nなどの基本条件の吟味がされるべきである。これらの点について、さらに詳しく述べる。

スペクトルの分類

スペクトルデータベースを有効に運用するためには、データの分類（ランキング）が必要となろう。もし、データベースのスペクトルを標準ピークとして用い、ピークフィティング法により定量分析を行うとすると、この標準ピークは感度が較正されているものでなければならぬし、また測定時のエネルギー分解能が確認されているものでなければならぬ。或は、XPSによる状態分析を行う場合は、ピークエネルギー値が較正されているものでなければならぬ。他方、それほど厳密に較正されていなくても、比較データとして参考にされればよい場合もある。一般に、厳しく較正を要求されたデータを測定することは大変であるから、データ提供者とデータ使用者の間に緩やかなインターフェースを設けることが現実的である。即ち、データにランキングを導入し、データ提供者の測定条件とデータ利用者の解析条件に柔軟に対応する事である。

ランクの提案：

Special Standard (SS)： 特別に設計された標準分光器で測定されたスペクトル。感度、エネルギー、分解能とも厳密に規定された特別標準スペクトル（分光器の較正などに使用する）。例えば、VAMASタイプのCMAで測定した後藤教授のスペクトルなど。

Absolute Standard (AS)： エネルギー軸と強度軸の両方が一定の水準以上の精度で較正されており、且つ較正方法がデータ中に明記されているスペクトル。

Energy Calibrated (EC)： エネルギー軸が一定の水準以上の精度で較正されており、且つ較正方法がデータ中に明記されているスペクトル。

Intensity Calibrated (IC)： 強度軸が一定の水準以上の精度で較正されており、且つ較正方法がデータ中に明記されているスペクトル。

Qualitative Reference (QR)： 特に較正を行っていない比較・参照用スペクトル。

SSクラス以外は、純元素の他に、化合物、混合物など応用的な試料のスペクトルも含む。

エネルギー軸の較正

エネルギー軸の較正はピークエネルギーの既知のピークを用いて行う。何人かの研究者によって基準となるピークの正確なエネルギー値が報告されている。⁵⁾ それらの

中でどれを使った方がよいとは断定できない。大切なことはどの値を用い、どの様な方法で較正したかを明記することである。そうすれば、後にさらに正確な値が報告されたとしても換算して再較正可能だからである。較正の方法についての説明はここでは省略するので文献を参照されたい。⁶⁾

装置関数の較正

測定されるスペクトルの強度は、測定条件や装置により異なる。他の装置で測定されたスペクトルを標準として定量解析を行う場合、この「装置による差」が問題となる。データを装置間で互換可能にするためには、装置関数の補正を行わねばならない。装置関数の補正には二つの要素がある。すなわち、強度と分解能である。これらは相互に影響しあうものだが、別々に求めて積をとり近似することも出来る。

強度の較正： いま、 $Q_1(E)$ 、 $Q_2(E)$ と装置関数の異なる二つの装置で同一励起ビーム強度で測定したときのスペクトルをそれぞれ $I_1(E)$ 、 $I_2(E)$ とする。二つのスペクトルの比をとると、

$$I_1(E) / I_2(E) = Q_1(E) / Q_2(E) \quad (1)$$

すなわち、スペクトル強度の比は装置関数比を与える。⁷⁾ 従って、共通の基準元素に対して、それぞれの装置でスペクトル測定を行い、装置関数比を求めておけば、相互換算が可能となる。基準元素としては、Cuが良いとされている。このため、ランクがASとICの標準スペクトルは、装置関数補正のため、その試料のスペクトルの他に、同一条件で測定したCuのN(E)スペクトルも添付する。この方法の詳細は文献を参照されたい。⁸⁾ この方法は現時点において相対的に可能性のある用法と思われるが、まだ完成されたものではないので、他の方法も入り込む余地を作っておく必要がある。

エネルギー分解能の補正： ピークフィティングなどにより定量を行う場合、被定量スペクトルと標準スペクトルとが同様のエネルギー分解能のものでなければならぬ。コンボリューション、デコンボリューションなどのテクニックを用いて分解能の補正を行うためには、それぞれのスペクトルの分解能が評価され既知で無ければならぬ。エネルギー分解能を評価する方法はいくつか提案されている。例を上げておく。

AESの場合：

- ① 2keVの弾性散乱ピークの幅で評価する方法
- ② Agのダブレットピーク高さの比で評価する方法
- ③ CuのP/B比で評価する方法
- ④ Cuのピーク幅で評価する方法

XPSの場合：

- ① Ag3d5/2のピーク幅で評価する方法
- ② Au4f7/2のピーク幅で評価する方法
- ③ Mo3d5/2のピーク幅で評価する方法

ランクがASとECのスペクトルは、分解能評価方法と結果を明記する。

S/N、感度の評価

S/Nはダメージなどの特別な事情が無い限り、良いほうが望ましい。S/Nの定義の仕方はいろいろ有るので、数値を評価するためには、定義のチェックが必須である。ここでは詳細は省略するので文献を参照されたい。⁹⁾

以上、スペクトルの質を押さえる上で重要な事項を述べたが、これは常に研究課題

なので、データベースの設計に当たって、将来提案される方法にも柔軟に対応できる入れ物を目指さねばならない。

6. ユーティリティソフト

スペクトルデータベースを利用する場合、色々なユーティリティソフトの助けが必要となろう。例えば、①自分のスペクトルとデータベースのスペクトルとがデータ測定間隔が一致しない場合、これらをどちらかに合わせるデータ変換ソフト、②装置関数補正のソフトウェア、③エネルギー分解能を一致させるためのコンボリユーション、デコボリユーションソフト、④各種ファイルフォーマット変換ソフト、などである。この他にも様々なソフトが必要になって来るであろう。データベースの構築と共に、ユーティリティソフトの開発を行わねばならない。

7. 運用

運用の方法については、これから十分議論されなければならない。いま考えられるのは次のような形態であろう。データの提供者は、管理者に電子メールまたはフロッピーディスクでデータを提出する。データの測定条件は自由である。ただし、条件はデータに添付されなければならない。同一の元素、化合物などについて複数のスペクトルが寄せられてきても構わない。むしろ歓迎である。選択は利用者が行えば良い。データが提出されたら、データベース委員会で評価する。要求される基準を満すことをチェックするためである。要求される基準のチェックとは、測定精度、校正精度の絶対値が高いか低いかではなく、方法が明記され検証性が高い否かである。校正方法自体の評価は変わるかも知れないからである。チェックにパスしたデータは管理者がデータベースへデータの登録を行う。データはネットワークのサーバに格納される。

データの供給の仕方はこれから更に検討を深めねばならない。通信で誰でも引き出せるのは問題である。しかし完ぺきに閉じてしまったら、多くのひとに使ってもらえない。できるだけ多くのひとに覗いてもらえるけれど、データの提供を受けるには、手数料の支払いや、会費の支払いを伴うような仕組みを研究することが要る。利用者は通信回線により検索し必要なデータのサブセットを選び出し、これらをデータベースより取り出す。データが多い場合は通信回線では転送に時間がかかるかもしれない。ユーザーへの大量のデータの輸送はフロッピーディスクやコンパクトディスクなどのほうが適す場合も考えられる。半年に1~2回の割合でデータを固定し、全データあるいは最新データをコンパクトディスクで供給することもアイディアである。運用に関してはまだ多くの議論を要する。

8. 著作権

著作権問題は次のように考えたらどうだろうか？ 個々のデータの著作権はデータの提供者に属し、データベースとしては表面分析研究会に属する。データの利用者には、データベース中のデータを用いる場合には、必ず引用を義務づけるよう明示する。

9. さいごに

今では、ほとんどの分析装置がコンピュータ化され、自動的に大量のデータが測定される様になった。データの解析についても、広くコンピュータが利用されている。今後、更に有効にコンピュータの利用を推し進めて行けば、測定装置の個性の違いや測定条件の違いを相互に換算し、測定データ間に互換性を持たせることも可能である。

また、誰でもピーク強度やピーク位置に対して、標準化されたソフトウェアによる解析で、同じデータから同じ結果を得ることができる。ピーク強度、ピーク位置が求められてから後の解析に於いても、コンピュータは重要な役割を果たす。例えば定量解析では、評価の確定されている計算アルゴリズム、信頼性の高い相対感度因子あるいは標準スペクトル、補正のための物理パラメータなどを準備しておけば、誰でも高い水準の定量解析が可能である。また、状態解析についても色々な物質についてのピーク位置データをいつでも照会できれば大変便利である。

以上述べたように、これからの表面分析の発展にとって、コンピュータを有効に利用するための基礎データの蓄積やソフトウェアの蓄積（広い意味でのデータベース）が極めて重要になるであろう。データベース利用により表面分析の精度、信頼性を向上させることが出来るし、またデータベースの利用を前提とすることによって初めて可能となる新しいデータ解析法も開発できるであろう。データベースは表面科学の進歩、材料評価技術向上に必須なものであろう。定義された方法で測定されたスペクトルの蓄積は、応用面だけでなく科学的にも重要な価値をもったものとなる。

最後に、「独創的なこととは最も普通のことを最も先に行なうことである」と認識している。この仕事には独創性を発揮できる可能性がある。従来、わが国ではこのような分野の貢献は少なく、外国のデータに依存することが多かった。我々もデータベース構築を通して科学の進歩に貢献したいと考えている。

文献

- 1) NIST X-ray Photoelectron Spectroscopy Database, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA
- 2) Sven Tougaard, QUASES Tougaard ApS., Ridderhatten 316, DK-5220 Odense S0, Denmark
- 3) 関根哲, B.V. Crist, 吉原一紘, ぶんせき, 11, 894 (1992)
- 4) 吉原一紘, 日本学術振興会マイクロビームアナリシス第81回研究会, 資料No.936, 12月, 72 (1994)
- 5) M.P.Seah, G.C.Smith, M.T. Anthony, NPL Report DMA (D), 691, February (1989)
- 6) D. Fujita and Y. Yoshihara, Surf. Interface Anal., 21, 226 (1994)
- 7) M.P.Seah, G.C.Smith, NPL Report DMA(D), 737, Sept.(1989)
- 8) 吉武道子, 吉原一紘, 表面科学, 15, 6, 376 (1994)
- 9) M.P.Seah and C.P. Hunt, NPL Report DMM (A), 101, April (1993)